

Roll No.....

Total No. of page(s): 1

BT-8/M-24: 48403-SE
PE-CS-AIDS- 416A Application of Data Science in Industry

Time: 3 hours]

[Max. Marks: 75

Note: Attempt Five questions in all, selecting at least one question from each unit.

Unit – 1

- Q. 1 a). Explain the terms Vectors and Matrices. 7.5 marks
b). How product of matrix and vector is done. Explain with the help of example. 7.5 marks
- Q. 2 How Eigen Value Decomposition is done. Explain. 15 marks

Unit – 2

- Q. 3 Explain the following:
a). Mean Variance
b). Multivariate normal distribution 15 marks
c). Covariance Matrix
- Q. 4 Illustrate the term Confidence Interval. When and how we use confidence interval for calculating for estimates. 15 marks

Unit – 3

- Q. 5. Explain the terms:
a). Simple Linear Regression
b). Multivariate Linear Regression 15 marks
- Q. 6 Write Algorithm for k-means clustering and explain it with the help of example. 15 marks

Unit – 4

- Q. 7 a). How Data Analysis problems can be solved using case study of Google Company. 7.5 marks
b). Write in brief case study of Walmart supply chain management. 7.5 marks
- Q. 8 a). What is Data Pipeline and how data pipeline can be made using airflow. 7.5 marks

BT-8/M-24

48403

APPLICATION OF DATA SCIENCE IN INDUSTRY

Paper-PE-CS-AIDS-416A

Time Allowed : 3 Hours]

[Maximum Marks : 75

Note : Attempt five questions in all, selecting at least one question from each Unit. All questions carry equal marks.

UNIT-I

1. (a) Explain the terms Rank and null space. 7½
(b) Explain the concept of vectors and projections in Geometric view. 7½
2. Explain with the help of example eigenvalue decomposition. 15

UNIT-II

3. Explain the following terms: 15
(a) Mean Variance (b) Covariance
(c) Descriptive statistics.
4. Explain in detail univariate and multivariate normal distributions. 15

UNIT-III

5. Explain the concept of k-means clustering algorithm with the help of example. 15
6. Illustrate the concept of Logistic Regression. Explain its working with the help of example. 15

UNIT-IV

7. Explain the case study of OLAP Tool for the food industry and Walmart supply chain management. 15
8. (a) How Real time data streaming with Apache Kafka with company data set can be done ? 7½
(b) How Data processing can be done using Spark streaming ? 7½